ESCO (2014) SEC 010 FINAL

Creation Date: 13/03/2014

Last update: 14/03/2014

Semantic interoperability with ESCO: pilot phase

Purpose of the document

The multilingual classification of European Skills, Competences, Qualifications and Occupations has the ambition to become a common language for the European labour market and the education/training sector. ESCO provides a common vocabulary that can bridge the communication gap between national classifications or thesauri. This will enable actors on the labour market and in education/training to exchange information – independent of classification systems or languages (semantic interoperability). Such a system of semantic interoperability is particularly relevant for EURES, the network of European public employment services (PES).[1]

Following the release of a pilot version of ESCO on 23 October 2013, semantic interoperability can now be tested in practice, as discussed by the ESCO Board. The ESCO Maintenance Committee and the ESCO Secretariat jointly developed an approach to test the interoperability potential of ESCO with selected national classifications. On the meeting of the Heads of Public Employment Services in December 2013 the Commission and the ESCO Board invited employment services to participate in this pilot phase.

In this document the ESCO Secretariat summarises the approach for testing the interoperability between ESCO and national classifications. The document clarifies the involvement that would be expected of participants and the expected outcomes.

[1] In January 2014 the Commission presented a proposal for a Regulation of the European Parliament and of the Council on a European network of Employment Services, workers' access to mobility services and the further integration of labour markets [COM (2014) 6 final]. One objective of this proposal is "to enable the EURES portal to carry out a good automated matching between job vacancies and CV's across Member States, translating in all EU languages and understanding skills, competences, occupations and qualifications acquired at national level".

The document will be presented to the ESCO Maintenance Committee, to the Assistants to the Heads of Public Employment Services, and to the contact persons in public employment services (PES) that expressed their interest in participating in the pilot phase.

Phased approach for creating mappings

The Commission services prepare for a phased implementation of semantic interoperability on the labour market:

- In a pilot phase PES and Commission services will test the creation of mappings between national occupational classifications and ESCO. The current pilot version of ESCO (ESCO v0, published on 23-24 October 2013) will be used for this pilot. The test results will feed into the development of a mapping environment (tools, documentation and training).
- Once the mapping environment is available, full scale mapping of all national occupational classifications can start. Process and planning for a full scale mapping will be defined at a later stage, taking into account the findings from the pilot phase.

Implementation of the pilot phase

The ESCO Secretariat suggests implementation of the interoperability testing in the following steps:

Step 1: Selecting employment services and classification systems

The European Commission invited public employment services during the meeting of the Heads of Public Employment Services in Vilnius (10/12/2013) to express their interest in participating in the pilot phase.

For the interoperability testing the ESCO Maintenance Committee and the ESCO Secretariat will select a number of classification systems managed by the participating PES. These will ideally include:

- *a very elaborated national occupational classification that is linked to a classification of skills and competences,*
- *a national occupational classification with a low level of detail,*
- *a national occupational classification based on ISCO,*
- *a national occupational classification that is not based on ISCO.*

Step 2: Creating mappings

In the course of the interoperability testing, a partial or full mapping (i.e. translation table) between the national occupational classification and ESCO will have to be created. This includes the following steps:

a) Transforming all selected classifications into the required data format (e.g. SKOS);
b) Importing them into tools for automated thesaurus alignment;
c) Assessing what level of quality can be achieved through automated mapping;
d) Finalising mappings for some or all economic sectors by reworking the mapping manually.

Resources needed from participating PES:

- Provide the national occupational classification in a machine readable format (e.g. RDF, XML, XLS);
- Participate in 2 technical workshops.

Expected outcomes:

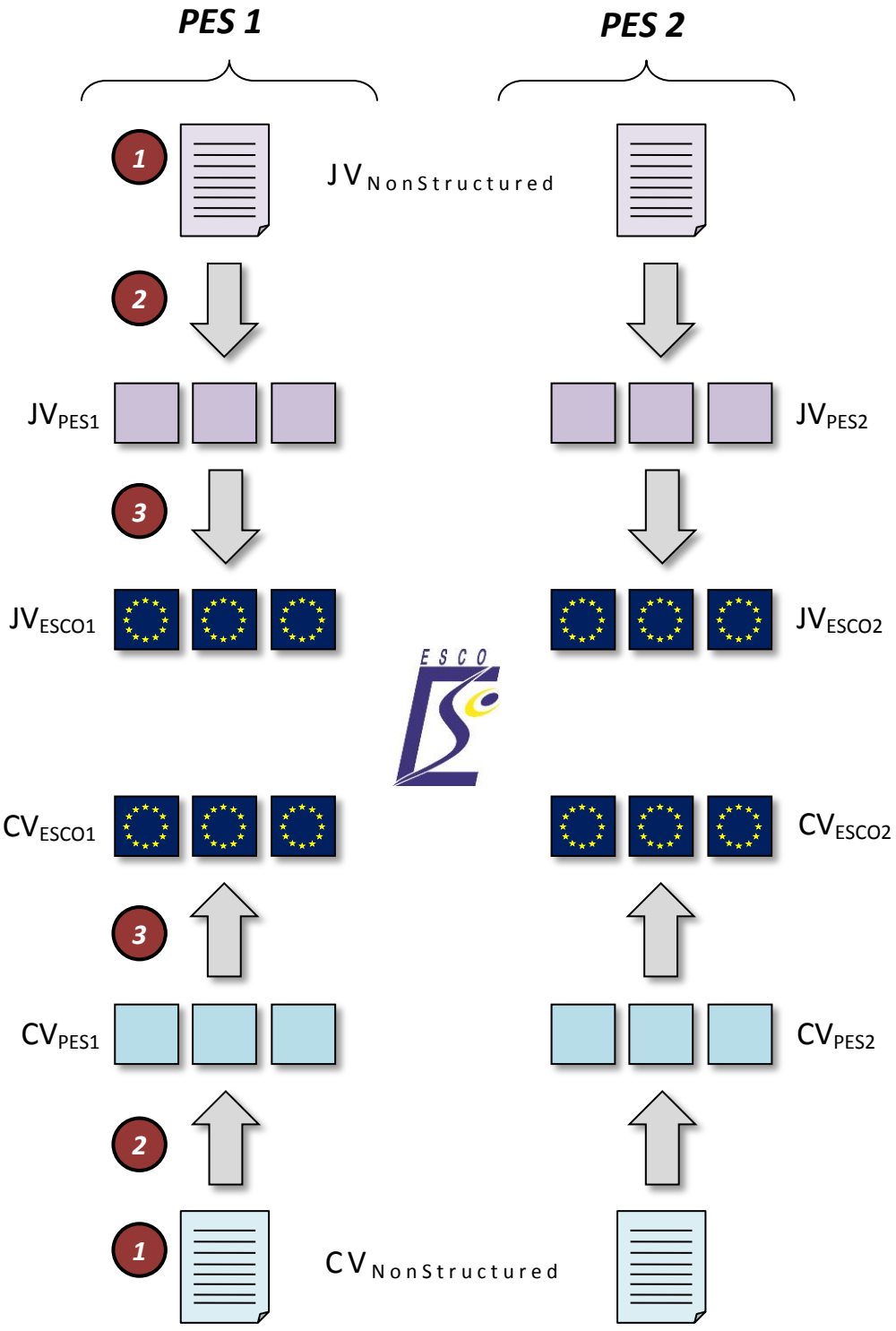
- Identify best practices for the creation of mappings;
- Assess effort/costs for creating and updating mappings in further detail;
- Understand the timings needed for implementation;
- Better understand which tools, documents and services can support the process and prioritise them.

Step 3: Assessing results

Once mappings have been created, participants in the pilot phase can assess their quality. Particular attention should be given to i) suitability of mappings for cross-border job matching, ii) suitability of mappings for skills-based job matching and iii) aspects of technical and organisational interoperability that might impact on semantic interoperability. The assessment will be done in four steps as illustrated in figure 1:

a) PES participating in the test will receive the *same* sample documents translated in the respective languages (1 in figure 1). These documents are delivered as unstructured data, i.e. plain text. Sample documents will be job vacancies ($JV_{NonStructured}$) as well as jobseekers' profiles or curricula vitae ($CV_{NonStructured}$). The sample data should include CVs and job vacancies that are expected to be full or partial matches in a job matching process.

Fig. 1: Approach for testing semantic interoperability between PES



b) Each participating PES transforms these documents into structured data according to their own procedures (2 in figure 1). The documents should be treated in a similar way as the organisation would treat a real job vacancy or CV that has been notified to the PES. The result of this process will be documents that are structured in the format as usually used in these PES (CV_{PES1}, JV_{PES1}, CV_{PES2} and JV_{PES2}). These documents will contain different types of data fields such as work experience, skills and competences, work environments, qualifications and many more. They will usually contain information encoded with the national classification system.

c) The structured information that has been produced in step 2 will be transformed into ESCO (3 in figure 1). The mapping created in step 2 will used for this purpose. The result are "Europeanised" variants of the CVs and job vacancies (CV_{ESCO1}, JV_{ESCO1}, CV_{ESCO2}, JV_{ESCO2})

d) After the stepwise transformation of documents into structured data and into ESCO data has been made, the results can be analysed in the following way:

i. Comparing $JV_{NonStructured} \Leftrightarrow JV_{PES1} \Leftrightarrow JV_{ESCO1}$ will allow better understanding how much information is lost in each step of transforming the data. A non-structured (and thus non-machine readable) job vacancy will contain more details than the structured data set in the PES. When "Europeanising" the data set by using ESCO, additional detail will be lost. The comparison of the three different data sets will allow assessing how much detail is lost in which step. CVs as well as documents of PES2 can be compared in the same manner.

ii. Comparing $JV_{ESCO1} \Leftrightarrow JV_{ESCO2}$ as well as $CV_{ESCO1} \Leftrightarrow CV_{ESCO2}$ will show if different processes for transforming non-structured data into ESCO data lead to similar results. In other words, it will reveal, if the two PES speak the same language by using ESCO or if the different application of ESCO will lead to divergence.

iii. Comparing $JV_{ESCO1} \Leftrightarrow CV_{ESCO2}$ as well as $JV_{ESCO2} \Leftrightarrow CV_{ESCO1}$ allows assessing if ESCO has the potential to enable skills-based job matching across borders. If a CV is chosen as sample data that partially matches a job vacancy, but where a certain set of skills/competences of the job seeker is missing, this information should be maintained when transforming the data into ESCO data. Suitability of the processed data for job-matching can be assessed manually or – if feasible – through testing in job matching algorithms.

Resources expected form participating PES:

- Process a small sample of client profiles and job vacancies according to their established processes;
- Participate in 3 technical workshops.

Expected outcomes:

- Assess what level of semantic interoperability can be reached by mapping to ESCO;
- Better understand and quantify potential benefits;
- Identify potential improvements of ESCO on a conceptual level;
- Understand what type of support national administrations need;
- Devise a methodology and understand in how far it needs to be adapted to specific classification systems.

Step 4: Quality improvement and validation (optional)

Employment services that would like to publish mappings (correspondence tables) as results of the testing exercise could improve the quality of the mapping in an iterative process. Once the quality of the mapping is validated by the employment service it can be published on the ESCO portal and prepared for reuse in machine-to-machine communication (publication as LOD[2], use in the ESCO API[3] and/or API for the PES). The decision whether to publish the mapping will be left to each participating employment service.

Expected outcomes:

- Publicly available mappings that can be used for further testing;
- Better understanding criteria for quality assessment and validation of mappings.

Step 5: Documentation and dissemination of results

A final report will summarise the results of the interoperability testing. It will be shared with all public employment services from countries participating in EURES or in the PROGRESS axis of EaSI, but also with other stakeholders that manage classification systems.

[2] Linked Open Data (LOD) "*describes a method of publishing structured data so that it can be interlinked and become more useful. […] Rather than using them to serve web pages for human readers, it extends them to share information in a way that can be read automatically by computers. This enables data from different sources to be connected and queried.*" (Wikipedia)

[3] "*An application programming interface (API) specifies how some software components should interact with each other.*" (Wikipedia)